



3 1761 11634745 1

10



Statistics
Canada

Statistique
Canada

CANSIM®

Services in
Canada



Statistics Canada

Statistics Canada is the country's largest producer and disseminator of statistical information. The agency provides information free or at minimum cost on almost every type of economic and social activity in Canada. Data are made available to the general public and to people in business, government, the media, labour organizations, and educational institutions.

In addition to some 1,300 publications of various periodicities issued every year, Statistics Canada also disseminates its information through such media as microfiche, microfilm, computer printouts, computer tapes and CANSIM.

CANSIM (Canadian Socio-Economic Information Management System)

This brochure describes the principal features of CANSIM, Statistics Canada's machine-readable database, and how it can be accessed. CANSIM comprises two modules, the CANSIM Time-Series module containing current and historical information from a broad range of interrelated socio-economic fields, and the CANSIM Cross-Classified module which addresses the demand for multi-dimensional data mostly originating from researchers in the fields of demography, education, health and justice.

CANSIM Time-Series

The Data

The CANSIM Time-Series module is often termed a "general purpose data bank" as it offers a detailed and conceptually compatible set of aggregate data pertaining to a large number of interrelated fields. The module is maintained by a single system through which data can be easily entered, stored, maintained and manipulated in a uniform general format.

These data are presented as time series which usually encompass the entire historical record available. The time series are grouped into larger blocks called "matrices". Thus the various series comprising the monthly Consumer Price Index for Canada, individual series that can be retrieved independently, are all grouped under a single "matrix" heading for easier identification and location in the base.

The CANSIM Time-Series module comprises major groups of data from the following subject-matter areas:

- System of National Accounts
- Prices and Price Indexes

- Labour
- Manufacturing and Primary Industries (including Fuel, Power and Mining)
- Capital and Finance
- Construction
- Merchandising and Services
- External Trade
- Transportation
- Agriculture and Food
- Population Estimates and Projections
- Health and Welfare

It further contains data from the *Canadian Statistical Review*, the *Bank of Canada Review* (including selected U.S. statistics) and the *Quebec Statistical Review*. The content is constantly growing as new sets of data are being added. Currently some 280 publications are represented in part or in full in the CANSIM Time-Series module.

Potential Users

Although access to the CANSIM system is not restricted to specific groups of statistical users, it is of special appeal to those individuals who use data regularly, either in analyses of a recurrent nature or in undertakings that require a historical perspective, such as trend analyses, projections or modelling. These include financial officers, corporate planners, investment and market analysts in medium and large size corporations, public and private sector economists, statisticians and researchers, as well as analysts in labour organizations.

Users will be drawn to CANSIM by the wide variety of analytical software available which permits quick and easy manipulation into forms more meaningful to the user. They will further appreciate the following features of the data base:

- daily updating and revision so that each data point in the historical record is the most current available in terms of the revision process;
- updated data are usually available on CANSIM at the same time as they are officially released through the "Statistics Canada Daily";
- all data entered are subject to edit checks during the entry operation to ensure accuracy and reliability;
- unpublished (but publishable) data can be made available;
- elimination of the costly and time-consuming problem of repetitively gathering, editing and putting data into machine-readable form;
- toll-free access from major Canadian cities via a type-writer terminal in either interactive or batch mode.

The CANSIM Time-Series module is offered to users in two ways: the CANSIM Main Base and the CANSIM Mini Base.

CANSIM Main Base

Definition

The CANSIM Main Base is the Statistics Canada system comprising the data themselves, the CANSIM data entry software, the CANSIM retrieval software and the CANSIM housekeeping software. It is the vehicle through which users can gain access to the full CANSIM Time-Series module.

The CANSIM Main Base currently contains some 300,000 time series and is updated on a daily basis, usually at the same time as the data are officially released. It is stored in on-line disk packs at the CANSIM host service bureau which is Datacrown Inc. (Systems Dimensions Limited at time of writing).

Access

The CANSIM retrieval software will extract the time series required from the CANSIM Main Base. This software is available in both interactive and batch mode for use from a typewriter terminal. Data can also be retrieved through a card-reader/printer.

The batch system will appeal to users who require large amounts of data, often on a tape or punched card format, for further processing as well as those who are involved in undertakings of a recurrent nature. Both will appreciate the cost reduction features of such a system. Manipulative programs are available for use with the data retrieved; these packages provide reporting, regression, modelling, forecasting and plotting facilities.

The CANSIM Interactive System (CIS) has been written in such a fashion as to be usable by individuals with little knowledge of computer software. No prior knowledge of APL, the language in which the CANSIM Interactive program is written, is necessary to accomplish CANSIM Main Base retrievals. CIS offers a choice of pre-programmed presentations when only data displays are desired; it will also make data available for further processing if so prompted.

Further CIS features are a number of special-purpose tables in a full interactive mode. Such tables are available in the fields of National Income and Expenditure Accounts, Key Economic Series, Labour Force, International Travel, Electrical Products Manufacturing, Census and others,

with more expected in the future. Textual files, containing an analysis of the latest Consumer Price Index for Canada and various cities, can also be retrieved.

Documentation

An up-to-date list of the data on the CANSIM Main Base is contained in a set of volumes organized so as to permit quick and easy identification and location of the required series. The *CANSIM Summary Reference Index* (Cat. No. 12-202E) summarizes the contents of the Main Base by subject and source publications. It further provides primary (Matrix) identification numbers which form the key to the *CANSIM Main Base Series Directory* (Cat. No. 12-203E). This Directory is a detailed guide to the data on the CANSIM Main Base; for each series it gives the title, start date, source of data, security level as well as the two identification numbers that can be used to retrieve series. It further identifies series that are no longer being updated, indicates which group or agency is responsible for the updating, revision and quality control of the series, and provides matrix notes informing the users of specific features of particular series in terms of collection of data or reference periods.

Information on access procedures to and data retrievals from the CANSIM Main Base are contained in the *CANSIM Interactive System Users Manual* (Cat. No. 12-553E) for access in interactive mode and in the *CANSIM Users Manual for Data Retrieval and Manipulation* (Cat. No. 12-531E) for access in batch mode or when using a card-reader/printer.

Documentation on access to the CANSIM host service bureau computer as well as on the software for time series manipulation in both interactive and batch mode is available from the CANSIM host service bureau and/or Statistics Canada.

CANSIM Mini Base

Definition

The CANSIM Mini Base is a standard subset of data drawn from the CANSIM Main Base and containing approximately 17,500 of the most widely used time series. The CANSIM Mini Base is updated daily, within the next working day of the series becoming available through the CANSIM Main Base.

The contents of the CANSIM Mini Base are revised twice a year, in March and September, on the basis of information about the use of series contained in the Main Base, the Mini Base and the Mini Base Supplement. All series in the Mini Base are public and none are terminated.

Access

The CANSIM Mini Base is offered by a number of Secondary Distributors. A Secondary Distributor is an organization (usually, but not always, a computer service bureau) that has agreed to disseminate CANSIM data according to Statistics Canada guidelines. Secondary Distributors provide the user with a large choice of software, telecommunications networks and technical support staff that can be utilized in accessing CANSIM data.

Besides the CANSIM Mini Base, Secondary Distributors can offer their users supplementary CANSIM series as part of their CANSIM Mini Base Supplement. This set of data will vary from one Secondary Distributor to another and enables each to meet unique data needs of their clients. Arrangements have been made whereby Secondary Distributors will receive historical series for their Mini Base Supplement within 24 hours of request and whereby supplementary series are updated the working day after those series are updated in the CANSIM Main Base, i.e. in exactly the same fashion as regular Mini Base series. This service is optional, however, and is not utilized by all Secondary Distributors. Statistics Canada also provides the capability for Secondary Distributors to produce documentation (Series Directory) for their Mini Base Supplement on an as needed basis.

The enclosed insert provides a current list of CANSIM Mini Base Secondary Distributors, including a brief description of their service, and a list of the Canadian cities from which they provide toll-free service.

Documentation

The *CANSIM Mini Base Series Directory* (Cat. No. 12-204E) is the Mini Base equivalent of the *CANSIM Summary Reference Index* and the *CANSIM Main Base Series Directory*. It provides the same information about its matrices and series as the latter does about the Main Base matrices and series, i.e. title, start date, source of data, identification numbers, responsible group or agency and matrix notes.

Information on the access and analytical software to the Mini Base and Mini Base Supplement must be requested from the Secondary Distributors.

CANSIM Cross-Classified

The System

Recognition of the increasing use of multi-dimensional information in machine-readable form has led to this

newest CANSIM module. The approach has been to adapt machine-processing techniques to current uses of multi-dimensional statistical aggregations.

Referencing multi-dimensional data is subject to wide variations depending on the type of aggregation and/or calculation required. For uniformity, statistical data are organized in this system according to a standard format — THE TABLE.

The structure of the cross-classified table allows data of up to nine levels of cross-classification to be entered, stored and retrieved. Descriptive information provides titles, footnotes, explanations, definitions and so on to foster proper use of the table.

Type of Data

Presently, priority is being given to loading groupings of social information collected by Statistics Canada.

Information on health, justice, education, science and culture has been identified. Tables are also being identified for 1976 Census data, agriculture, and family income and expenditure. These early data sets on the cross-classified base are considered a set of general interest statistics which will have a recurring interest to users, but they are only a fraction of the information that can potentially be made available.

From a technical point of view there are virtually no limitations to the level of detail that can be included.

Access

Interactive Processing

The cross-classified system features two conversational functions. They provide a set of commands that are used to retrieve and display descriptive information and to manipulate the data base. Using simple words to construct a command string, the user can:

- display part or all of the descriptive texts for the various tables;
- display footnotes;
- retrieve and display subsets of tables;
- load table data and/or user defined subsets or create APL variables which can then be used with APL Public Libraries or user defined functions for manipulation.

Command String

DIS 00010101 O = L P = '1964'/'1974' D(1) = CAN, QUE, ONT, BC D(2) = M, F

This command will retrieve an array of data which can be described as the number of live births in Canada, Quebec, Ontario and British Columbia for males and females from 1964 to 1974. These data can be displayed and/or manipulated in APL. Once a table has been identified on the base, any combination of dimensions and periods can be retrieved.

XCL Utility

To support data manipulation, this software has been developed to allow users to reformat retrieved data so that they can be used with APL Public Libraries. The user needs no knowledge of APL. Data can be plotted, histograms can be created and basic descriptive statistics automatically generated.

Table Producing Language*

This is a computer system designed to select, restructure, cross-tabulate and display statistical data. The syntax of the language allows the user to construct expressions that define the required table. English-like phrases are used which eliminate the need for knowledge of any programming language. Tables can be retrieved and displayed either at a terminal or via remote, high-speed printers. Data can be retrieved from one table or can be retrieved from a number of tables to create a new output table.

Table Statement

TABLE DIVORCE: SEX BY AGE, NO BY PERIOD;

TPL statements instruct the computer as to which data are to be tabulated and how they should be formatted. To improve table layout capabilities, additional statements are available to alter the format of the requested table, e.g., titles and additional footnotes can be added; page and column widths can be changed; stubs and headers can be relabelled, etc.

Potential Users

The CANSIM Cross-Classified Module will be of particular interest to analysts and researchers specializing in social studies and societal evolution and problems. It meets the needs for cross-tabulation and analysis of data on social conditions (family structure, occupation, income and others) available from the Census with data on social phenomena in the fields of education, health, welfare, justice and others.

*Developed by the Division of General Systems of the Bureau of Labor Statistics, U.S. Department of Labor, Washington, D.C.

Service

The CANSIM Staff

The CANSIM staff, in addition to being responsible for entering, maintaining, revising and updating the data in both modules, will assist users with a wide variety of questions and problems in the use of CANSIM. They will answer inquiries about timeliness, text and data accuracy, secure series and demand for a new data on the Main Base, as well as about the CANSIM system.

They will also perform retrievals from the CANSIM Main Base for users who are not equipped with a terminal or who require output on cards or tapes. The charges for such services are broken down into four categories:

- Less than 1,000 series:
 - 15¢ per series:
 - minimum of \$5.00 for a computer printout
 - minimum of \$25.00 for outputs on tape (user supplied)
- 1,000 series or more:
 - computer cost plus 50% (any output)
- Manipulated data:
 - 50¢ per operation required
 - minimum of \$5.00
- Standing requests (the same service at regular intervals):
 - computer cost plus 50% (any output). Note: This is considerably less expensive than 15¢ per series.

Cross-classified Base — Direct Access

Batch retrieval, depending on the format of the required table, will vary from a minimum of \$9.00 to a maximum of \$20.00 per 1000 data points. These costs include system access, data retrieval and table structuring. It should be noted that these costs will not double with the second 1000 points. Rather, the costs will increase slightly up to the first 10,000 data points. Additional cost information is available on request.

Most CANSIM documents must be ordered through the CANSIM staff who must be contacted by potential users to obtain authorization to access the CANSIM Main Base.

The CANSIM staff can be reached at:

CANSIM
Statistics Canada
R.H. Coats Building, 23rd Floor
Tunney's Pasture
Ottawa, Ontario K1A 0Z8
(613) 995-7406

User Advisory Services

User Advisory Services is a division of Statistics Canada whose primary role is to help people identify, locate, obtain and use Statistics Canada data effectively. Trained staff provide advice on the availability, meaning and use of the data.

User Advisory Services has offices in nine major cities across the country. Toll-free telephone lines to the offices are available in the Maritimes and throughout Saskatchewan and Alberta.

Each local office has one or more Regional Advisors who assist both experienced and inexperienced users with their data problems. They promote the use of statistical data, including CANSIM, through participation in conferences, convention displays, workshops, seminars and briefing sessions so as to introduce Statistics Canada data to existing and prospective users. They can help people get in touch with subject-matter specialists in Ottawa if necessary, and they provide feedback to Statistics Canada data-producing divisions on the statistical needs and problems of users so as to help the agency improve the content and availability of its products and services.

Each local office is equipped with a low speed terminal to access the CANSIM Main Base. CANSIM retrievals can thus be performed, at cost, for the benefit of regional users without terminals. Information on CANSIM can be provided to users and potential users, more specifically, officers are able to answer inquiries with regard to use of CANSIM documentation, series identification, underlying methods and concepts on data series, requests for new data and so on.

Across Canada, User Advisory Services offices can be reached at:

St. John's,
Newfoundland
Tel: 726-0713

Halifax,
Nova Scotia
Tel: 426-5331

Montréal, Québec
Tel: 283-5725

Ottawa, Ontario
Tel: 992-4734

Toronto, Ontario
Tel: 966-6586

Winnipeg, Manitoba
Tel: 949-4020

Regina,
Saskatchewan
Tel: 569-5405

Edmonton, Alberta
Tel: 425-5052

Vancouver,
British Columbia
Tel: 666-3695

Also, toll-free access is provided in Nova Scotia, New Brunswick and Prince Edward Island by telephoning 1-800-565-7192. Throughout Saskatchewan, the Regina office can be reached by dialing 1-800-667-3524. Throughout Alberta, the Edmonton office can be reached by dialing 1-800-222-6400.

Summary: How to Access CANSIM

Users wishing to access the CANSIM system have a variety of options available to them. The choice of the method of access will depend on the user's needs and priorities.

Time Series Module

With respect to the Time Series module, the options available to the user are:

- to access the Main Base (or the Mini Base) directly via computer terminal through Datacrown Inc. Users may also arrange for special runs to be done by Datacrown, if they do not have their own terminal.
- to access the Mini Base directly via computer terminal through any one of the CANSIM Secondary Distributors listed on the insert. The Secondary Distributor may add supplementary series from the Main Base (in addition to those series contained in the Mini Base) to the data they make available to serve their customers. Users without terminals may order special runs from Secondary Distributors.
- users without terminals may order special runs on any of the data contained on the Main Base from the CANSIM staff. Such requests may be on a regular or occasional basis and will be cost-recoverable.
- users may get a limited number of runs on an occasional basis from any of the local offices of User Advisory Services.

Cross-Sectional Module

At the present time, users interested in the Cross-classified module may:

- access the system directly via computer terminal through Datacrown Inc. (Systems Dimensions Limited at time of writing). Users may also arrange for special runs to be done by Datacrown, if they do not have their own terminal.
- order special runs to be done by the CANSIM staff. Such requests may be on a regular or occasional basis and will be cost-recoverable.
- obtain a limited number of runs on an occasional basis from any of the Regional Offices of User Advisory Services.

On peut aussi communiquer sans frais en Nouvelle-

Écosse, au Nouveau-Brunswick et à l'Île-du-Prince-

Édouard en composant le 1-800-565-7192. En Saskatchewan-

man, on peut rejoindre le bureau de Regina en composant le 780-667-3524 et en Alberta, le bureau d'Edmonton, au

1-800-222-6400.

Sommaire: comment consulter CANSIM

L'utilisateur qui désire consulter la banque de données CANSIM dispose de plusieurs options. Il doit choisir sa méthode d'accès en fonction de ses besoins et de ses préférences.

Module des séries chronologiques

Ce module offre aux utilisateurs les options suivantes:

- Consulter la Base principale (ou la Mini Base) directement au moyen du réseau de terminaux de la société Datacrow Inc. (présentement Systems Dimensions Limited) et faire des passages spéciaux par cette Base. S'ils n'ont pas leur propre terminal, ils peuvent aussi commander des passages spéciaux de leurs clients. Les utilisateurs qui n'ont pas de terminal peuvent pas d'un terminal, demander au personnel de l'Assistance-utilisateurs de faire des passages spéciaux à partir des données de la Base principale et faire faire, à l'occasion, un nombre limité de passages par l'un des bureaux locaux de l'Assistance-utilisateurs.

Module des tableaux de recoupement

À l'heure actuelle, les utilisateurs qui s'intéressent aux tableaux de recoupement disposent des possibilités suivantes:

- Consulter la banque de données directement au moyen du réseau de terminaux de la société Datacrow Inc. (présentement Systems Dimensions Limited) et faire des passages spéciaux par cette société, s'ils n'ont pas leur propre terminal.
- Commander, régulièrement ou occasionnellement, au personnel de CANSIM de faire des passages spéciaux, à condition d'en assumer les frais.
- Faire faire, à l'occasion, un nombre limité de passages spéciaux par l'un des bureaux locaux de l'Assistance-utilisateurs.

Il faut s'adresser à:

CANSIM
Statistique Canada
Imm. R.H. Coats, 23^e étage
Parc Tunney
Ottawa, Ont. K1A 0Z8
(613) 995-7406

Assistance-utilisateurs

L'Assistance-utilisateurs est une division de Statistique Canada dont le rôle premier est d'aider les gens à reconnaître, à repérer, à obtenir et à utiliser les données de Statistique Canada de manière efficace. Un personnel qualifié fournit des conseils concernant la disponibilité, le sens et l'utilisation des données.

L'Assistance-utilisateurs a des bureaux dans neuf grandes villes du pays, ainsi qu'un service téléphonique gratuit dans les Maritimes, en Saskatchewan et en Alberta.

Chaque bureau local compte au moins un conseiller régional qui aide les utilisateurs, expérimentés ou non, à résoudre leurs problèmes. Ces conseillers encouragent les gens à se servir des données statistiques, notamment de CANSIM, en participant à des conférences, des stands de congrès, des ateliers de travail, des séminaires et des séances d'information. Au besoin, ils peuvent les mettre en communication avec des spécialistes d'Ottawa. Ils informent aussi les divisions produisant des données à Statistique Canada des besoins et des problèmes statistiques des utilisateurs afin que le Bureau puisse améliorer le contenu et la disponibilité de ses produits et de ses services.

Chaque bureau local dispose d'un terminal à faible débit pour consulter la Base principale CANSIM. Les utilisateurs qui n'ont pas de terminal peuvent ainsi obtenir des données CANSIM au prix coûtant. Les conseillers des bureaux locaux peuvent donner de l'information sur CANSIM à tous les utilisateurs actuels et éventuels, notamment en répondant à leurs questions concernant l'emploi de la documentation CANSIM, l'identification des séries, les méthodes et concepts à la base des séries, les demandes de nouvelles données et bien d'autres sujets. On peut communiquer avec l'Assistance-utilisateurs en composant l'un des numéros suivants:

Saint-Jean (T.-N.) 726-0713
Winnipeg (Manitoba) 949-4020

Halifax (N.-E.) 426-5331
Montréal (Québec) 283-5725

Ottawa (Ontario) 996-6586
Edmonton (Alberta) 425-5052

Vancouver (C.-B.) 666-3695
Toronto (Ontario) 966-6586

ées sur les conditions sociales (structure des familles, occupation, revenu, etc.) tirées du recensement et, en part, des données sur les phénomènes sociaux dans les domaines de l'éducation, de la santé, du bien-être, de la justice, et autres.

Personnel CANSIM

L'être responsable de l'entrée, de l'entretien, de la mise à jour des données dans les deux fichiers CANSIM aide les utilisateurs en leur fournissant des renseignements sur l'actualité des données et des chiffres des données et en tentant de résoudre divers problèmes. Il doit répondre à des demandes concernant les données introduites dans la Base principale, le système CANSIM dans son ensemble. Le personnel se charge aussi d'extraire des données de la Base principale CANSIM pour les utilisateurs qui n'ont pas le terminal ou qui désirent obtenir des sorties sur des bandes. Les frais sont répartis en quatre séries: 1.000 séries: minimum de \$5 pour un imprimé d'ordinateur minimum de \$25 pour les données sorties sur bande (fournie par l'utilisateur) 2. séries et plus: minimum de traitement plus 50% (toutes présentations) 3. données manipulées: 50¢ par opération 4. minimum de \$5 pour les données permanentes (même service demandé à des utilisateurs réguliers): Il est à noter que ceci est beaucoup moins cher que le tarif de 15¢ la série.

Accès direct à la base de données recueillies

L'extraction par lots est facturée en fonction de la présentation du tableau et varie d'un minimum de \$9 à un maximum de \$20 pour 1000 données simples. Les frais commencent à partir de \$20 pour 1000 données simples et augmentent légèrement jusqu'à 1000 données simples. On peut demander plus d'information au sujet des frais.

La plupart des documents CANSIM doivent être commandés par l'intermédiaire du personnel CANSIM, auquel les utilisateurs doivent aussi avoir recours pour obtenir l'auto-édition de consulter la Base principale CANSIM.

Suite d'instruction
DIS 00010101 0 = L P = '1964'/'1974' D(1) = CAN, QUE,
ONT, BC D(2) = M, F

Cette instruction a pour fonction de retrouver une série de données représentant le nombre de naissances vivantes au Canada, au Québec, en Ontario et en Colombie-Britannique, pour les hommes et les femmes, de 1964 à 1974. Ces données peuvent être manipulées et présentées en APL. Une fois le tableau repéré dans la base, toute combinaison de dimensions et de périodes peut être extraite.

Programme utilitaire XCL

Ce logiciel a été créé pour faciliter la manipulation des données, en permettant aux utilisateurs de restructurer les données extraites de façon qu'elles puissent être employées avec les bibliothèques publiques APL. L'utilisateur n'a pas besoin de connaître le langage APL. Il peut présenter des données au moyen d'un tracé, dessiner des histogrammes ou produire automatiquement des statistiques descriptives de base.

Langage de production de tableaux (TPL)*

Le TPL est un système informatique conçu pour choisir, réorganiser, recouper et afficher des données statistiques. La syntaxe de ce langage permet à l'utilisateur de construire des expressions qui définissent les tableaux requis. Comme les termes employés ressemblent à l'anglais, il n'est pas nécessaire de connaître un langage de programmation. Les tableaux peuvent être extraits et présentés au moyen d'un terminal ou d'une imprimante rapide à distance. Quant aux données, on peut les extraire d'un ou plusieurs tableaux pour en créer un nouveau.

L'énoncé Table

TABLE DIVORCE: SEXE BY AGE, NO BY PERIOD;

Utilisateurs éventuels

Les instructions TPL indiquent à l'ordinateur quelles données doivent être mises sous forme de tableau et comment elles doivent être disposées. Pour accroître les options offertes, on a prévu des instructions permettant de modifier la présentation d'un tableau: on peut ajouter des titres ou des renvois en bas de page, modifier la largeur des pages et des colonnes, changer les intitulés de lignes et de colonnes, etc.

Le module des tableaux de recoupement CANSIM sera particulièrement utile aux analystes et chercheurs qui s'intéressent aux études sociales, aux problèmes et à l'évolution de la société. Il répond à la demande d'analyses et de tableaux de recoupement établis d'une part, à partir de

*Créé par la Division des systèmes généraux du "Bureau of Labor Statistics, U.S. Department of Labor", Washington, D.C.

Tableaux de recoupement CANSIM

Le système

Devant l'accroissement de la demande pour des informations multi-dimensionnelles et ordinolinguées, nous avons créé ce dernier module de CANSIM. Nous avons essayé d'adapter les techniques de traitement par machine aux tendances qui caractérisent actuellement l'utilisation d'agréats statistiques multi-dimensionnels. L'établissement d'un index de données est susceptible de divergences considérables selon le genre d'agrégation et de calcul requis. Pour en assurer l'uniformité, les statistiques de ce système sont présentées sous une forme normalisée, le TABLEAU.

La structure d'un tableau de recoupement permet d'enregistrer, de stocker et d'extraire un maximum de neuf séries qui se recoupent. Des renseignements écrits dans les titres, notes, explications, définitions, etc., qui favorisent la bonne utilisation du tableau.

Genre de données

Pour l'instant, nous travaillons surtout au chargement de blocs de données à caractère social recueillies par Statistique Canada. Nous avons procédé à la définition de renseignements sur la santé, la justice, l'éducation, les sciences et la culture. Nous sommes aussi en train de définir des tableaux relatifs au recensement de 1976, à l'agriculture et aux revenus et dépenses des familles. Ces premiers ensembles constituent en quelque sorte une introduction à la base recoupée. D'intérêt général et d'actualité durable pour les utilisateurs, ils ne représentent qu'une fraction de l'information qu'il est possible d'obtenir.

Du point de vue technique, il n'y a presque aucune limite à la finesse des séries.

Accès

Traitement en mode interactif

Le système de recoupement possède deux fonctions conversationnelles comportant toute une série d'instructions qui servent à extraire et à présenter des renseignements descriptifs et à manipuler la base de données. En consultant une suite d'instructions à l'aide de mots simples, l'utilisateur peut:

- présenter une partie ou l'ensemble des textes descriptifs pour les divers tableaux;
- présenter les notes connexes;
- extraire et présenter des sous-ensembles;
- charger les données d'un tableau ou les sous-ensembles qu'il a définis ou encore créer des variables en APL qui peuvent ensuite être utilisées de concert avec les bibliothèques publiques APL ou les fonctions définies par lui pour des fins de manipulation.

Mini Base et du supplément de la Mini Base. Toutes les séries de la Mini Base sont publiques et aucunes ne sont définitives.

Accès

On peut consulter la Mini Base CANSIM par l'intermédiaire d'un certain nombre de distributeurs secondaires. Un distributeur secondaire est un organisme (habituellement, mais pas nécessairement, une société de services informatiques) qui a consenti à diffuser les données CANSIM conformément aux principes directeurs formulés par Statistique Canada. Ces distributeurs offrent à l'utilisateur toute une gamme de logiciels, de réseaux de télécommunications, ainsi que les services d'un personnel de soutien technique, ce qui facilite l'accès à CANSIM.

Outre la Mini Base CANSIM, les distributeurs secondaires peuvent aussi fournir à leurs utilisateurs des séries supplémentaires tirées du supplément de la Mini Base CANSIM. Cet ensemble de données varie d'un distributeur à l'autre, permettant à chacun de répondre aux besoins particuliers de ses clients. Des dispositions ont été prises pour que les distributeurs secondaires reçoivent les séries chronologiques supplémentaires dans un délai de 24 heures après en avoir fait la demande, et pour que les séries supplémentaires soient mises à jour une journée ouvrable après leur mise à jour dans la Base principale, tout comme les séries ordinaires de la Mini Base. Il convient toutefois de souligner que ce service est optionnel et n'est pas utilisé par tous les distributeurs secondaires. Statistique Canada offre aussi aux distributeurs la possibilité de produire les documents (répertoire des séries) nécessaires à leur supplément de la Mini Base.

Vous trouverez ci-joint une liste à jour des distributeurs secondaires de la Mini Base CANSIM, ainsi qu'une brève description de leurs services et une liste des villes canadiennes à partir desquelles ils fournissent un service gratuit.

Documentation

La publication intitulée *Mini Base CANSIM, répertoire des séries* (n° 12-204F au catalogue) est l'équivalent, pour la Mini Base, du *Répertoire de concordance des matrices CANSIM* et de *Base principale CANSIM, répertoire des séries*. Elle fait état des mêmes renseignements concernant la matrice et les séries, c'est-à-dire le titre, la date de départ, la source des données, le numéro d'identification, le groupe ou l'organisme responsable, et les notes relatives aux matrices.

Les distributeurs secondaires sont seuls à donner de l'information concernant l'accès et le logiciel analytique pour la Mini Base et son supplément.

active, des voyages internationaux, de la fabrication de produits électriques, du recensement, et l'on prévoit que beaucoup d'autres s'y ajouteront à l'avenir. Il est aussi possible d'extraire des dossiers textuels, renfermant des analyses du plus récent indice des prix à la consommation au Canada et dans diverses grandes villes.

Documentation

Il existe une série de publications qui donnent la liste, mise à jour, des données que renferme la Base principale CANSIM; elles ont été conçues pour que les utilisateurs puissent retrouver facilement et rapidement la série qu'ils recherchent. Le *Répertoire de concordance des matrices CANSIM* (n° 12-202F au catalogue) résume le contenu de la Base principale selon le sujet et la source, tout en indiquant les numéros d'identification (matriciels) primaires qui sont la clé de la *Base principale CANSIM, répertoire des séries* (n° 12-203F au catalogue). Ce dernier est un guide détaillé des données de la base principale CANSIM; on y trouve, pour chacune des séries, le titre, la date de départ, la source des données, la cote de sécurité et les deux numéros d'identification qui peuvent servir à extraire la série. Il indique aussi si les séries ont cessé d'être mises à jour, mentionne le groupe ou l'organisme responsable de la mise à jour, de la révision et du contrôle de la qualité et, dans des notes matricielles, signale aux utilisateurs certaines caractéristiques particulières à une série donnée (méthode de collecte, période de référence, etc.).

Pour des informations sur les méthodes d'accès aux données et d'extraction de ces données de la Base principale CANSIM, il faut consulter les publications suivantes: *Système interaction, manuel de l'utilisateur* (n° 12-553F) pour ce qui est du mode interactif, et *CANSIM: Manuel d'extraction et de manipulation des données* (n° 12-531F) pour l'accès par lots ou l'emploi d'un combine lecteur de cartes-imprimante.

La société informatique responsable du service CANSIM et/ou Statistique Canada offre de la documentation sur l'accès à son ordinateur et sur le logiciel servant aux manipulations de séries chronologiques en mode interactif ou par lots.

Mini base CANSIM

Définition

La Mini Base CANSIM est un sous-ensemble normalisé de données tirées de la Base principale CANSIM. Actuellement, elle contient environ 17,500 des séries chronologiques les plus utilisées. Les séries font l'objet d'une mise à jour quotidienne dans un délai d'une journée ouvrable après leur mise en disponibilité dans la Base principale CANSIM.

Le contenu de la Mini Base est révisé deux fois l'an, en mars et en septembre, à l'aide de renseignements concernant l'utilisation des séries de la Base principale, de la

moyen d'un terminal à clavier fonctionnant en mode interactif ou par lots.

Le module des séries chronologiques CANSIM se présente sous deux formes: la Base principale et la Mini Base CANSIM.

Base principale CANSIM

Définition

La Base principale CANSIM est le système de Statistique Canada qui se compose de données, du logiciel d'entrée des données, du logiciel d'extraction et du logiciel de servitude CANSIM. Elle permet aux utilisateurs d'avoir accès à tout le module des séries chronologiques CANSIM.

La Base principale CANSIM se compose actuellement d'environ 300,000 séries chronologiques qui sont mises à jour quotidiennement, généralement lorsque les données sont diffusées officiellement. Elle se trouve entreposée dans des chargeurs de disques en liaison directe avec l'ordinateur de la société qui assure les services informatiques CANSIM, soit Datacrown Inc. (présentement Systems Dimensions Limited).

Accès

Le logiciel d'extraction CANSIM peut extraire les séries chronologiques demandées de la Base principale CANSIM. Le logiciel, exploité par terminal à clavier, existe en mode interactif et par lots. On peut aussi obtenir les données grâce à un combiné lecteur de cartes-imprimante.

Le traitement par lots intéresse généralement les utilisateurs qui font des travaux de nature périodique ou encore ceux qui demandent un volume important de données, souvent sous forme de bandes ou de cartes perforées, en vue d'autres traitements. Les deux groupes d'utilisateurs seront heureux d'apprendre que ce genre de système réduit le coût des travaux. Ils pourront aussi obtenir des programmes de manipulation à employer avec les données extraites; ces logiciels permettent de faire des comptes rendus, des régressions, de la modélisation, de la prévision et des tracés.

Le système interaction CANSIM (SIC) a été conçu pour des gens qui connaissent peu le logiciel d'ordinateur. Pour extraire des données de la Base principale CANSIM, il n'est pas nécessaire de connaître le langage APL, le langage de programmation utilisé par le système interaction CANSIM. Le SIC offre tout un éventail de programmes pré-établis servant simplement à afficher des données, et fournit aussi, sur demande, des données destinées à des traitements ultérieurs.

Le SIC peut aussi produire en mode interactif un certain nombre de tableaux spécialisés, notamment dans les domaines des comptes nationaux des revenus et des dépenses, des séries économiques clés, de la population

- Travail
- Industries manufacturières et primaires (Combustibles, énergie et mines compris)
- Capitaux et finances
- Construction
- Commerce et services
- Commerce extérieur
- Transports
- Aliments et agriculture
- Estimations et projections démographiques
- Santé et bien-être social

Ce module renferme aussi des données tirées de la *Revue statistique du Canada* de la *Revue de la Banque du Canada* (y compris certaines statistiques américaines) et de la *Revue statistique du Québec*. Son contenu ne cesse de s'accroître à mesure que de nouveaux ensembles de données s'y ajoutent. Actuellement, quelque 280 publications sont intégrées en tout ou en partie au module des séries chronologiques de CANSIM.

Utilisateurs éventuels

L'accès au système CANSIM n'est pas restreint à des groupes particuliers d'utilisateurs de statistiques, mais il intéresse surtout les personnes qui se servent régulièrement de données, soit pour faire des analyses de nature périodique, soit pour des travaux exigeant une perspective historique, comme des analyses de tendances, des projections ou des modélisations. Les utilisateurs peuvent donc être des agents financiers, des planificateurs, des analystes d'investissements et de marchés travaillant pour la moyenne ou la grande entreprise, des économistes, des statisticiens et des chercheurs du secteur public ou privé, ou encore des analystes oeuvrant dans des organisations syndicales.

Toutes ces personnes peuvent s'intéresser à CANSIM car il fournit tout un éventail de logiciel analytique qui permet de manipuler des données facilement et rapidement. Il offre aussi les avantages suivants:

- les renseignements sont mis à jour et révisés quotidiennement, de sorte que toutes les données simples des dossiers d'archives sont toujours les plus actuelles possible;
- les données mises à jour sont habituellement disponibles sur CANSIM dès qu'elles ont été diffusées officiellement par l'entremise du "Quotidien" de Statistique Canada;
- toutes les données sont soumises à une certaine vérification à l'entrée afin d'assurer l'exactitude et la fiabilité;
- les données inédites (mais publiables) peuvent être obtenues;
- on élimine ainsi les travaux onéreux et répétitifs de collecte, de vérification et de mise sous forme ordonnée qui sont si coûteux;
- les utilisateurs peuvent obtenir les données gratuitement dans la plupart des grandes villes du Canada, au

Statistique Canada est le plus grand producteur et diffuseur de données statistiques au pays. L'organisme fournit, gratuitement ou contre une somme minimale, de l'information sur presque tous les genres d'activités économiques et sociales au Canada. Toutes les données sont mises à la disposition du grand public et des personnes travaillant dans les établissements d'enseignement, les médias, les organisations syndicales, les entreprises et l'administration publique.

Outre les quelque 1,300 bulletins publiés à intervalles divers au cours de l'année, Statistique Canada diffuse aussi de l'information par l'intermédiaire de microfiches, de microfilms, d'imprimés et de bandes d'ordinateur, ainsi qu'au moyen de CANSIM.

CANSIM (Système Canadien d'information socio-économique)

Ce feuillet décrit les principales caractéristiques de CANSIM, la base de données ordinales de Statistique Canada ainsi que les mécanismes pour y avoir accès. CANSIM se compose de deux modules: le module des séries chronologiques, qui renferme des données actuelles et historiques sur un grand nombre de sujets socio-économiques, et le module des tableaux de recoupement, qui répond à la demande de données multi-dimensionnelles provenant surtout de chercheurs des domaines de la démographie, de l'éducation, de la santé et de la justice.

Séries chronologiques CANSIM

Les données

Le module des séries chronologiques CANSIM est souvent considéré comme une "banque de données polyvalente", car il offre un ensemble de données détaillé, cohérent sur le plan des concepts et portant sur un éventail de données connexes. Ce module est mis à jour par un seul système, grâce auquel les données sont facilement entrées, stockées et manipulées sous une forme générale homogène.

Ces séries chronologiques englobent généralement tous les dossiers d'archives existants. Elles sont regroupées en ensembles plus grands, nommés "matrices". Ainsi, les diverses séries comprenant l'indice mensuel des prix à la consommation au Canada, qui constituent des séries distinctes et peuvent être extraites indépendamment, sont toutes regroupées dans une seule "matrice" pour qu'il soit plus facile de les localiser dans la base de données. Le module des séries chronologiques CANSIM comprend des séries de données sur les sujets suivants:

- Système de la comptabilité nationale
- Prix et indices de prix

CAI
BS
-Z310

Statistique
Canada



Statistics
Canada

Services
CANSIM®
au Canada